Computational modeling of RNA-protein binding interactions under an external force

Ralf Bundschuh

Department of Physics
Department of Chemistry and Biochemistry
Division of Hematology, Department of Internal Medicine
Center for RNA Biology
Interdisciplinary Biophysics Graduate Program

Ohio State University

July 7, 2026

Overview

Overview

RNA protein interactions

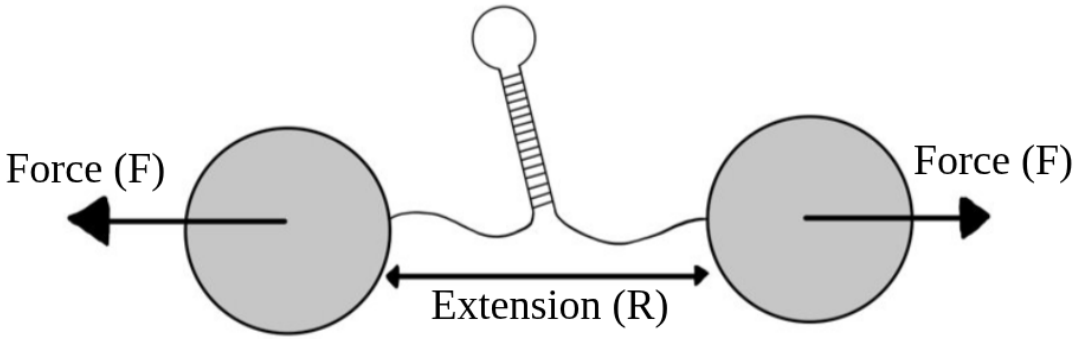
RNA protein interactions

- RNA by itself is awesome
- However, in cells RNA often interacts with many other players including proteins
- E.g., 500-2000 of human proteins are RNA binding
- Several methods to study these interactions
  - EMSA's for binding affinity
  - proximity labeling for yes/no
  - CLIP for binding sites
  - X-ray, cryo-EM, NMR for structures
- Here: Could force spectroscopy single molecule experiments tell us something about RNA protein interactions?

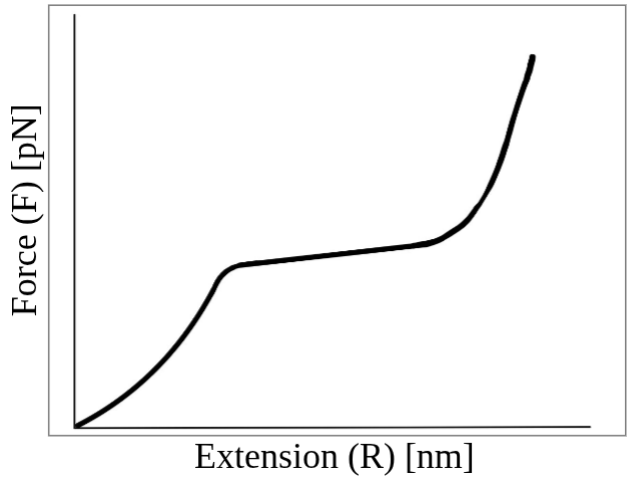
Why might force spectroscopy experiments work?

Force spectroscopy

RNA force spectroscopy concept



Force extension curve (FEC)



Connection to proteins

- Proteins have structural preferences (e.g., single-stranded RNA)

$\Rightarrow$ Proteins change the structural ensemble of an RNA

$\Rightarrow$ Changes in end-to-end distance ensemble

$\Rightarrow$ Changes in FECs

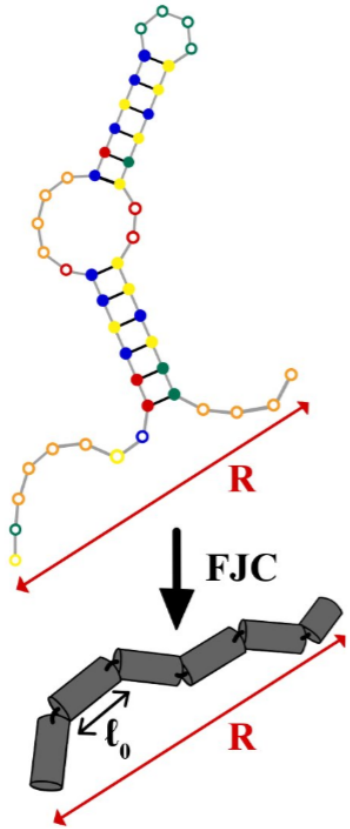
Overview

Partition function modeling of RNA secondary structure

- Structures s
- Energy of each structure $\Delta G(s)$
- Partition function $Z = \sum_s e^{-\frac{\Delta G(s)}{k_B T}}$
- Can be computed in $O(N^3)$ per McCaskill (we use Vienna implementation)
- Ensemble free energy $G = -k_B T \ln Z$

Partition function modeling of FECs

RNA as Freely Jointed Chain



Partition function

- Single segment of FJC

$$Z_{rod}(F) = \frac{k_B T}{\ell_0 F} \sinh\left(\frac{\ell_0 F}{k_B T}\right)$$

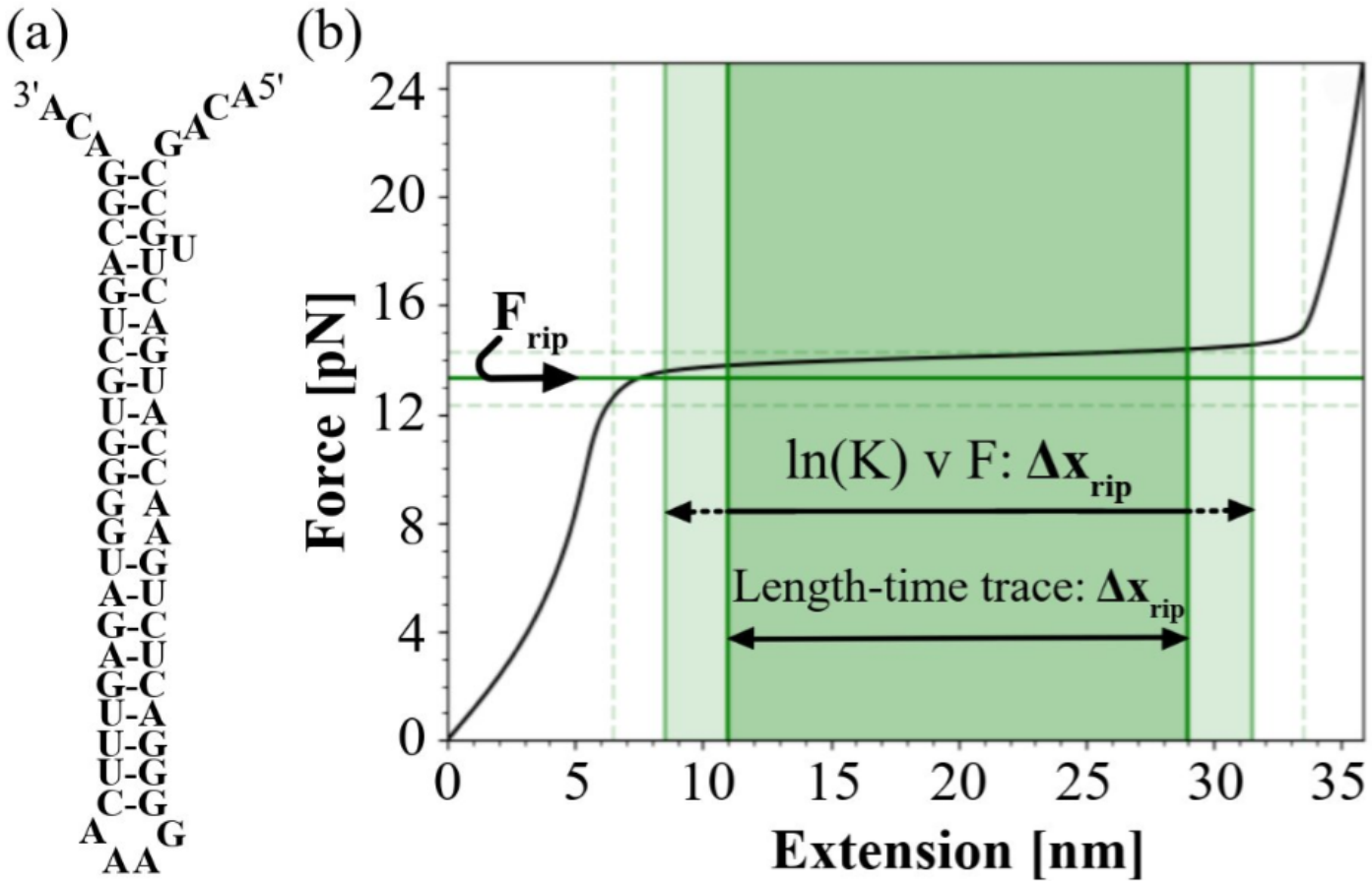
- Free energy
  $G(F) = -k_B T \ln Z_{rod}(F)$
- FEC: $R = -\frac{\mathrm{d}G}{\mathrm{d}F}$

For the whole structure ensemble:

$$Z(F) = \sum_s [Z_{rod}(F)]^{N_{seg}(s)} e^{-\frac{\Delta G(s)}{k_B T}}$$

Making sure it works

P5ab hairpin - comparison to Bustamante et al. experiments



Overview

Modeling RNA protein interactions

One binding site

- Simplest case: RNA with one binding site for protein
- Need affinity K_D
- Then (Hackermüller *et al.*),

$$Z(c) = \sum_s \left(1 + \frac{c}{K_D}\delta_{site,s}\right) e^{-\frac{\Delta G(s)}{k_B T}}$$

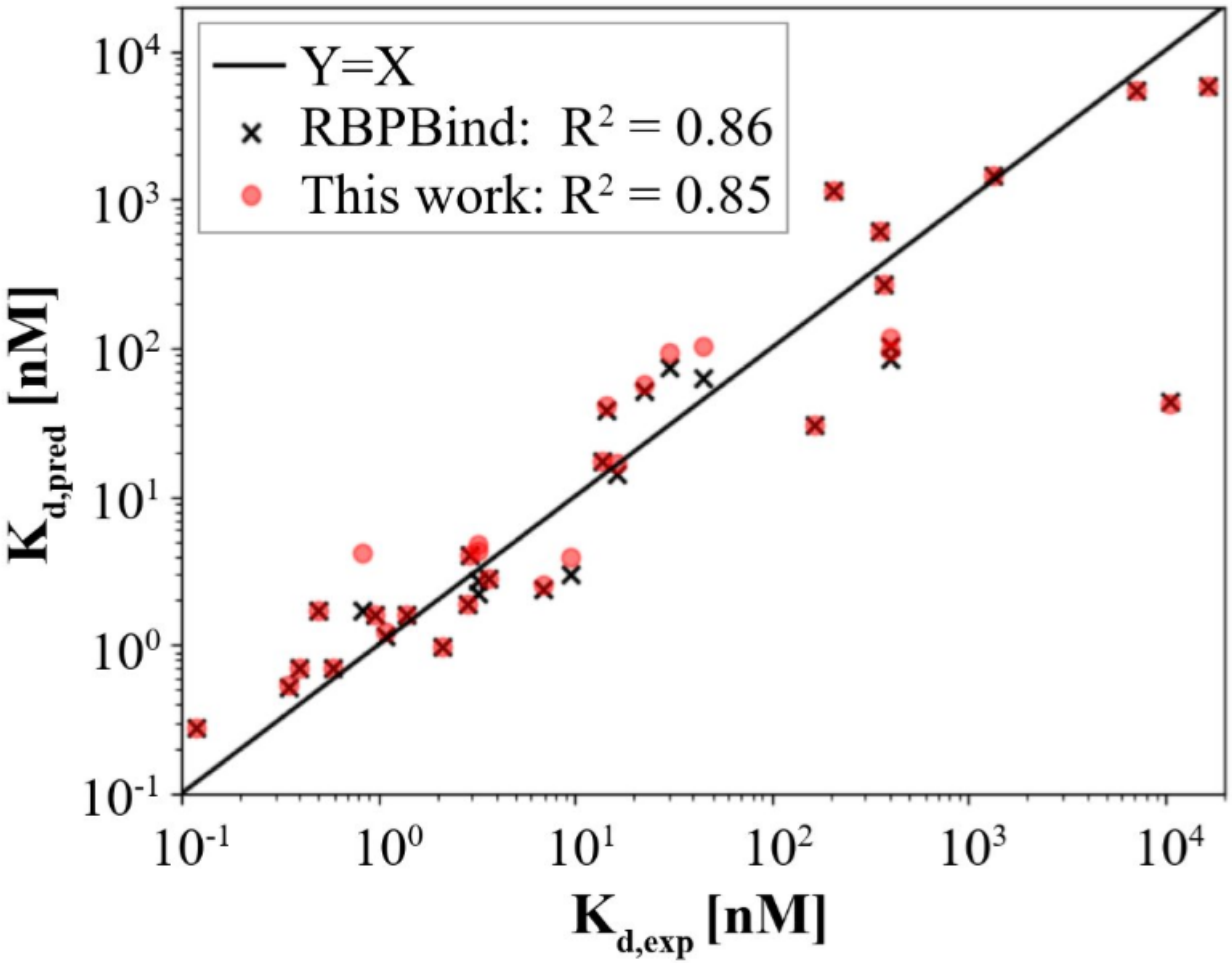
Arbitray binding sites

- Get table of binding affinities for each k-mer from RNAcompete experiments (footprint size k)
- Then,

$$Z(c) = \sum_s \left[\prod_{i=1}^{N-k+1} \left(1 + \frac{c}{K_{D,i}}\delta_{site_i,s}\right)\right] e^{-\frac{\Delta G(s)}{k_B T}}$$

Making sure it works

Comparison to biochemical experiments on HuR, RBFOX1, U2AF2, and KHDRBS3



Overview

Combining forces and proteins

Forces and proteins

- Force alone:

$$Z(F) = \sum_s [Z_{rod}(F)]^{N_{seg}(s)} e^{-\frac{\Delta G(s)}{k_B T}}$$

- Protein alone:

$$Z(c) = \sum_s \left[\prod_{i=1}^{N-k+1} \left(1 + \frac{c}{K_{D,i}} \delta_{site_i,s} \right) \right] e^{-\frac{\Delta G(s)}{k_B T}}$$

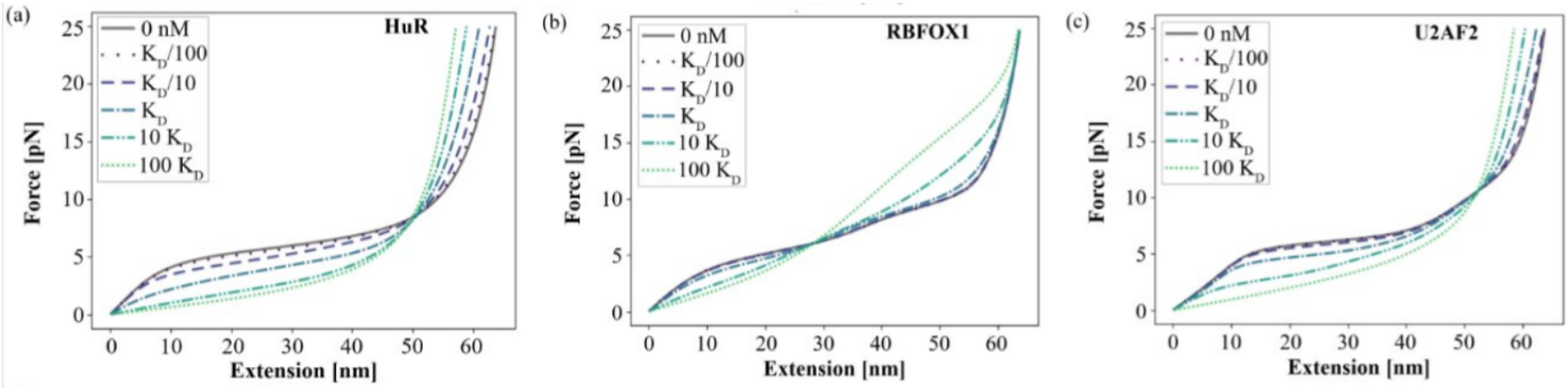
- Both together:

$$\begin{aligned} Z(c,F) &= \sum_s [Z_{FEC-rod}(F)]^{N_{seg}(s)} e^{-\frac{\Delta G(s)}{k_B T}} \\ &\times \prod_{i=1}^{N-k+1} \left(1 + \chi_{protein-rod-ext}(F) \frac{c}{K_{D,i}} \delta_{site_i,s} \right) \end{aligned}$$

Example FECs

Computing FECs for different protein concentrations

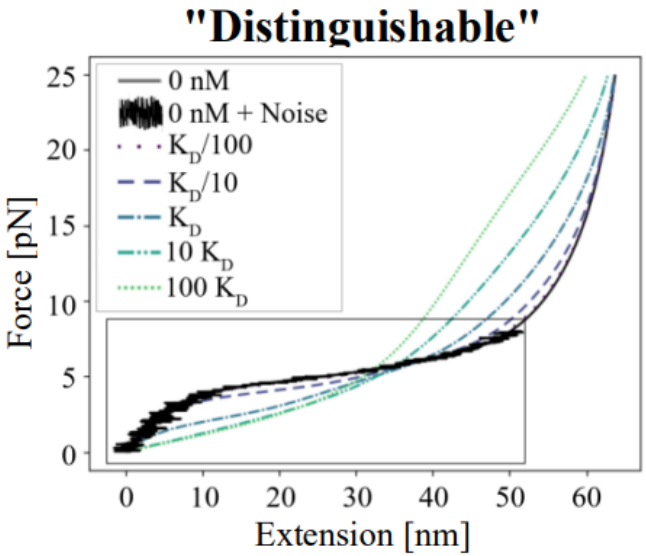
- Select 100nt sequences around natural binding sites of proteins
- Determine computational K_D (at zero force)
- Calculate $Z(c, F)$ for different c and F
- Calculate FECs at different protein concentrations

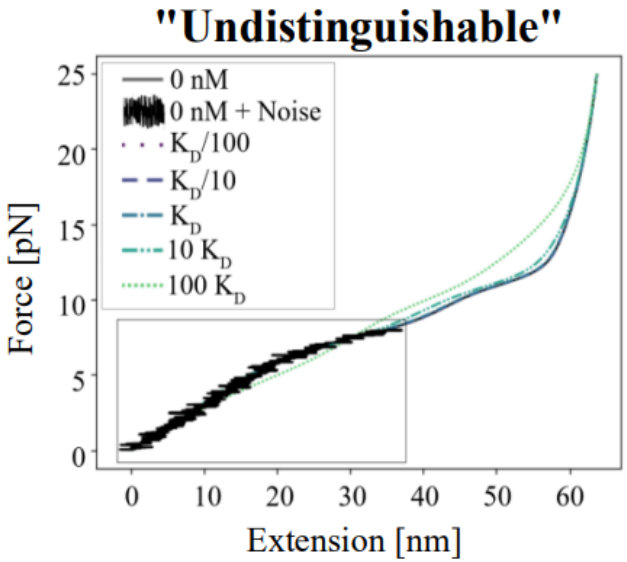


Can protein binding be distinguished?

Is protein binding distinguishable?

- Add noise according to experiments to $c = 0$ FEC
- Calculate average and std. dev. of area under curve up to F_{max}
- Vary c and calculate area under curve up to F_{max}
- Distinguishable if area under the curve at least 2σ from $c = 0$ area

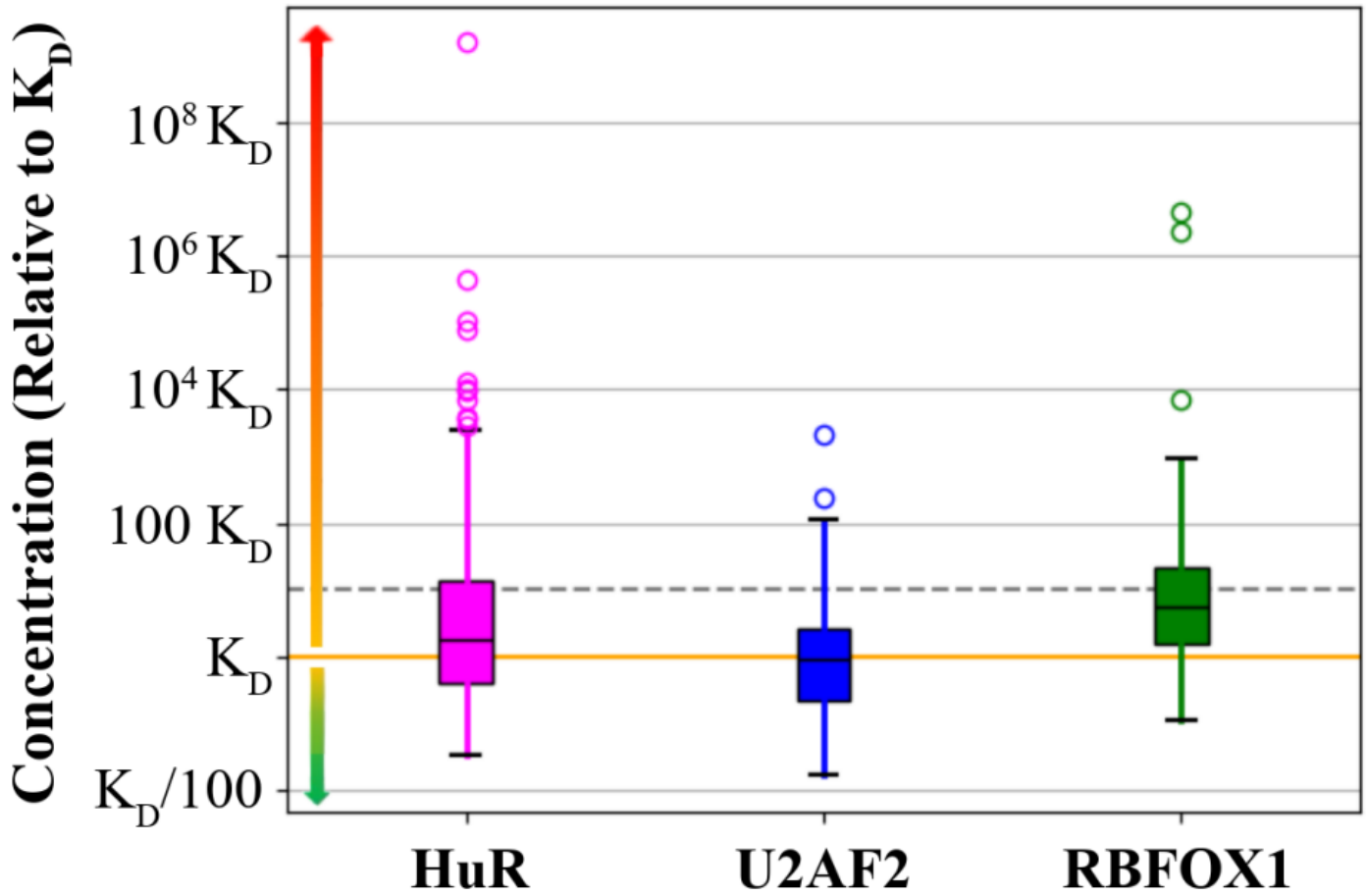




Becomes distinguishable at high enough concentrations for nearly every sequence!

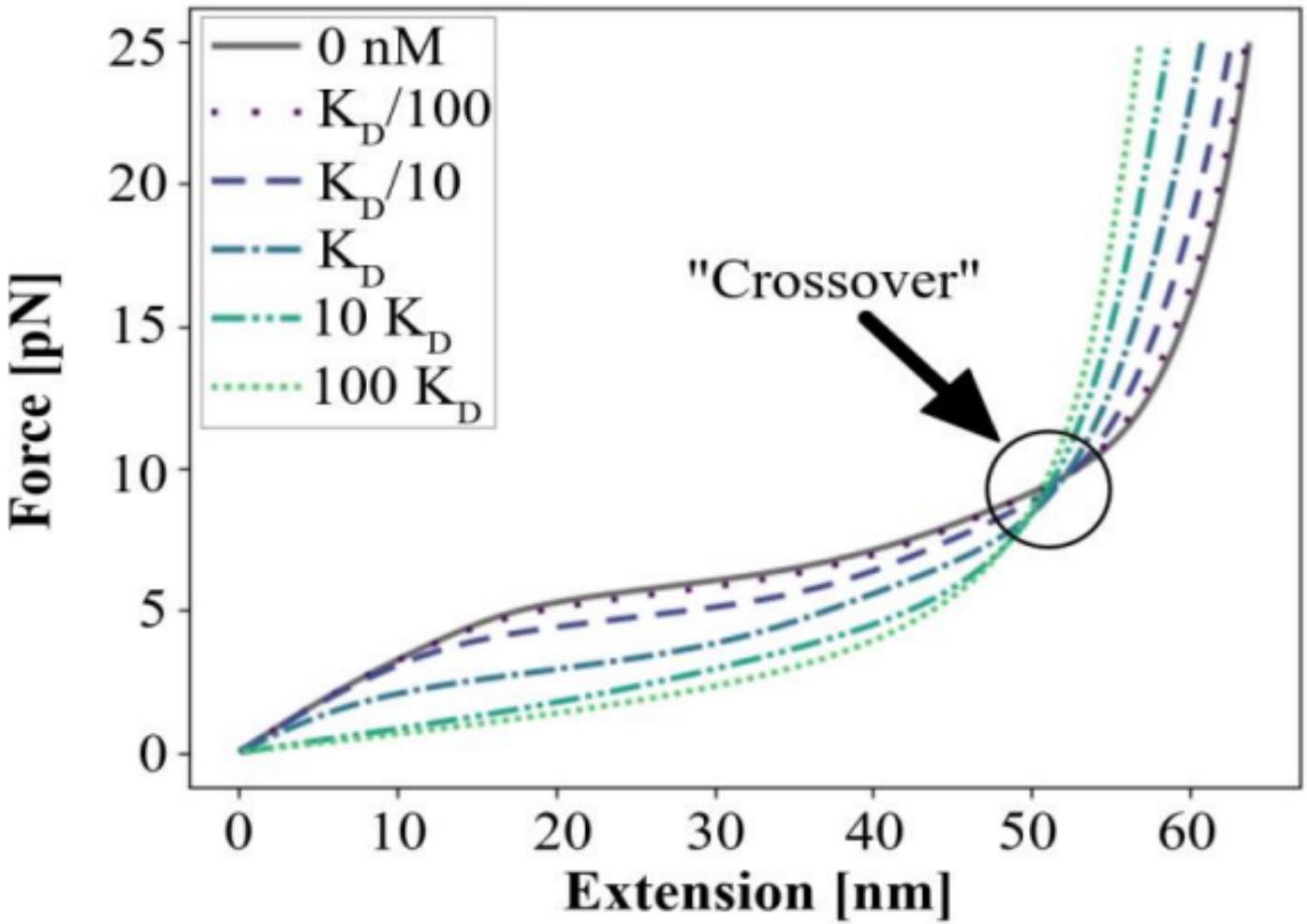
Can protein binding be distinguished at reasonable concentrations?

Is protein binding distinguishable at reasonable concentrations?



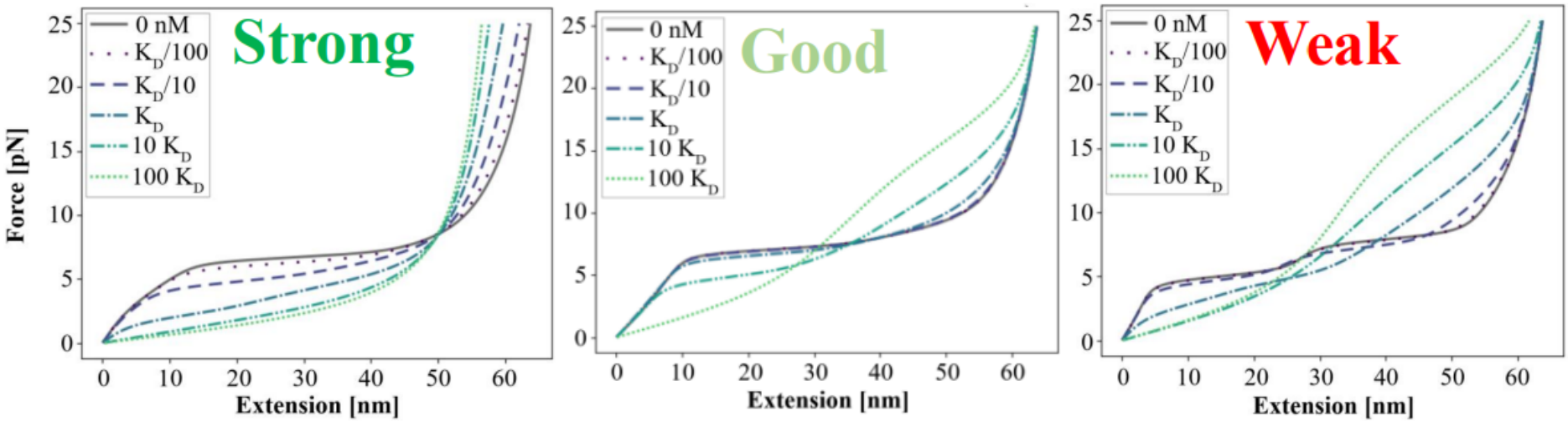
Crossover points

Notable feature: crossover points



Does every sequence have crossover points?

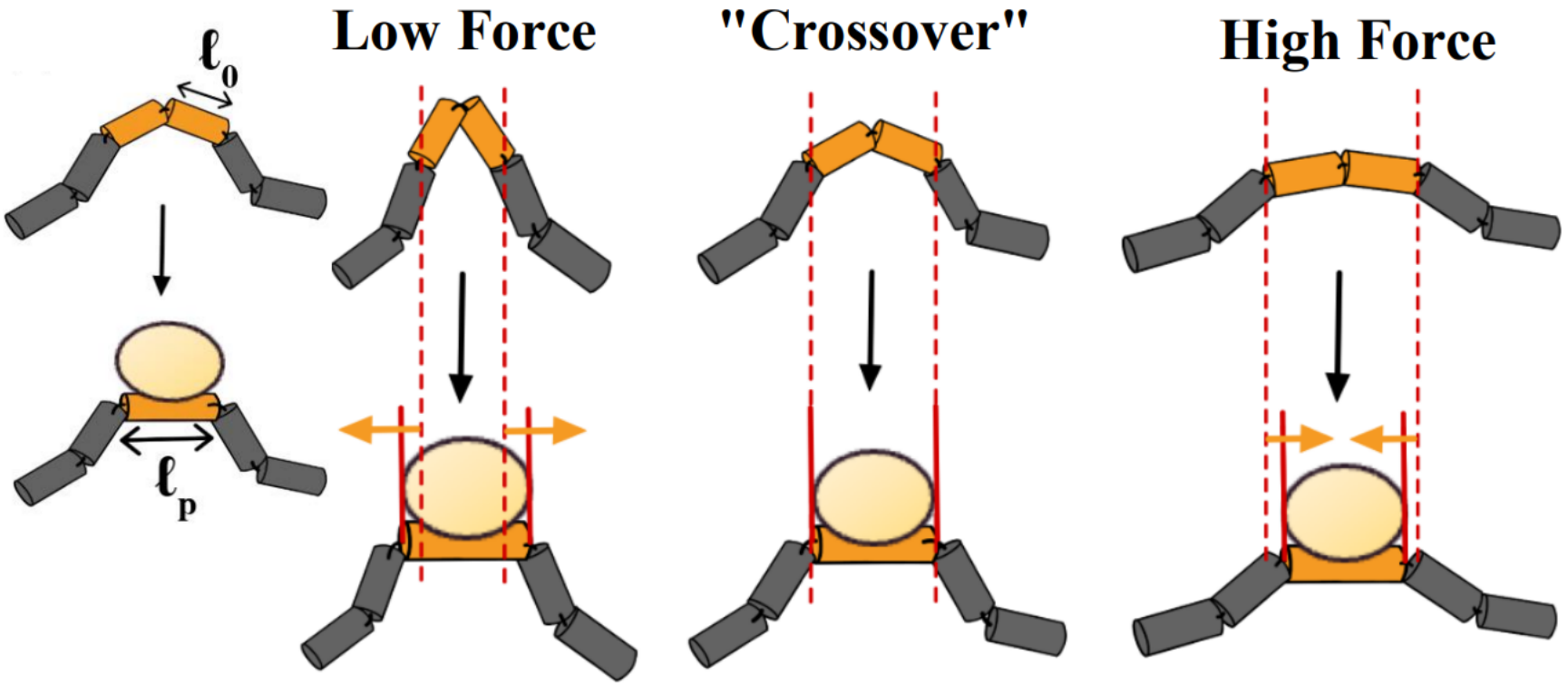
Define strength of crossover points



- Quantify through standard deviation of crossing forces
- 85% of sequences show "good" or better crossover points

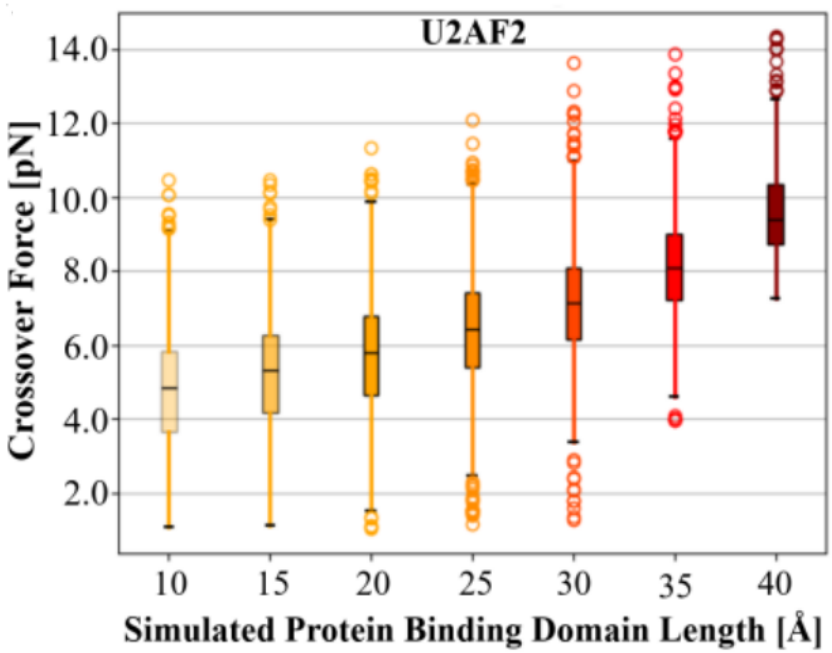
Intuition for crossover points

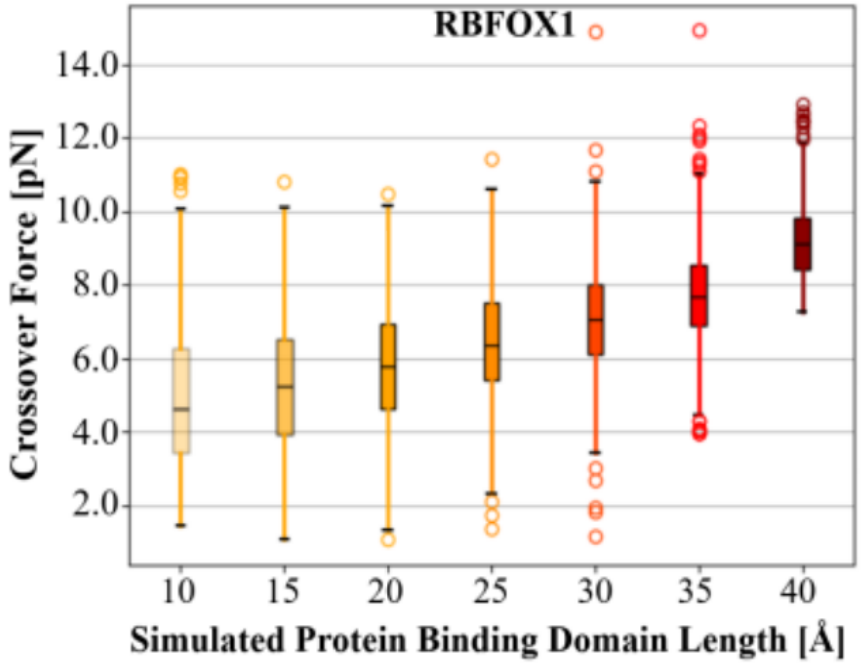
Probable mechanism for crossover points



Protein domain length dependence

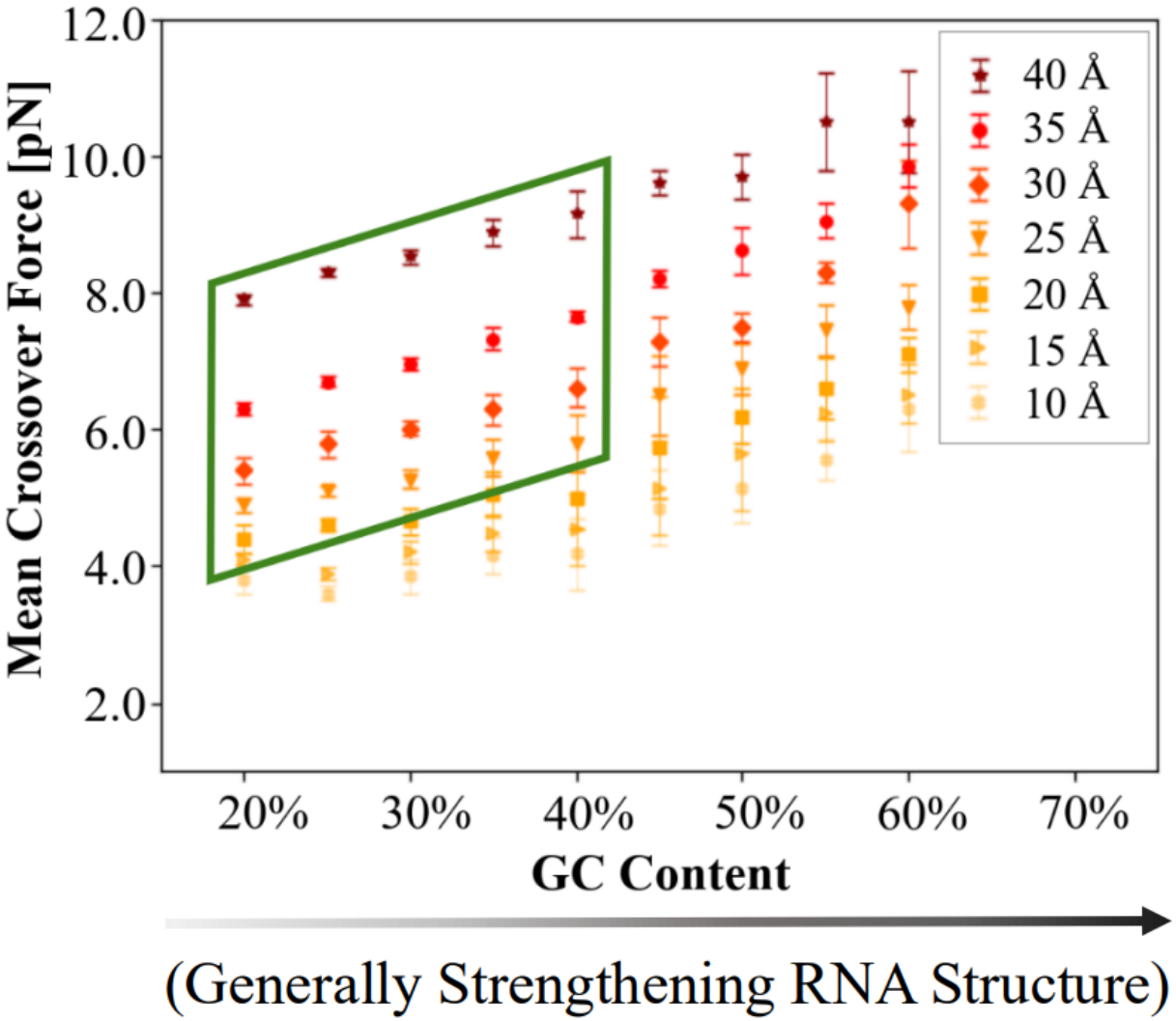
If mechanism is correct, crossover point should depend on protein binding geometry





Geometry from crossover points

Then, one should be able to extract geometry from cross-over points



Overview

Conclusions and Outlook

Conclusions

- Single molecule force spectroscopy experiments can measure RNA-RBP binding interactions at relevant concentrations
- Concentration-dependent force extension curves exhibit important crossover point feature
- Geometry of protein binding site affects force of crossover points
- Length of protein binding domain is extractable

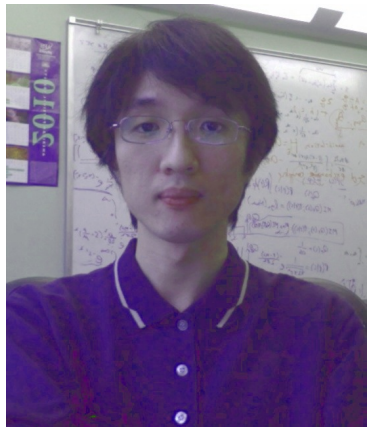
Future work

- Double stranded RNA binding proteins (dsRBPs)
- Try to convince an experimentalist to do these experiments
- RNA-RBP kinetics modeling

Who actually did the work

Danielle Wampler

Bob Forties



Yi-Hsuan Lin

Jeff Gaither